

101534764
1000000000000000

BUNDESREPUBLIK DEUTSCHLAND

EP03/12660



RECEIVED
23 MAR 2004
WIPO PCT

Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

Aktenzeichen: 102 52 457.2

Anmeldetag: 12. November 2002

Anmelder/Inhaber: Harman Becker Automotive Systems GmbH,
76307 Karlsbad/DE

Erstanmelder: DaimlerChrysler AG,
70567 Stuttgart/DE

Bezeichnung: Spracheingabe-Interface

IPC: G 10 L 15/22

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 05. März 2004
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag

Hintermeier

A 9161
06/00
EDV-L

BEST AVAILABLE COPY

DaimlerChrysler AG

Die/Fi

Spracheingabe-Interface

Die Erfindung betrifft Systeme, bei denen das Sprechen eines Anwenders aufgenommen und als Signal weitergeleitet wird.
5 Insbesondere betrifft die Erfindung Systeme, bei denen über eine Spracheingabe Funktionen ausgelöst oder gesteuert werden.

Systeme mit Spracheingabe und Sprachsteuerung sind bekannt.
10 So sind beispielsweise bei Computern Eingaben - Diktieren von Texten oder Steuerbefehle - per Sprache möglich. In der Medizintechnik existieren Geräte, die zur Entlastung bei schwierigen Eingriffen durch Sprachkommandos des Arztes gesteuert werden können. Auch im Sicherheitsbereich werden Vorrichtungen eingesetzt, die beispielsweise eine verschlossene Tür selektiv nur auf Spracheingaben autorisierter Personen öffnen.

Alle derartigen System erfordern für das Erkennen der sprachlichen Information bzw. Kommandos eine hohe akustische Qualität der Spracheingabe. Besonders störend wirkt sich beispielsweise ein zu großer Abstand zwischen Mikrofon und Mund des Sprechers aus (zu geringe Eingabelautstärke). In gleicher Weise wirkt ein Sprechen in eine Richtung, die nicht im Hauptaufnahmebereich des Mikrofons liegt. Auch eine relativ nahe Distanz zum Mikrofon kann sich negativ auswirken, da in diesem Fall zum einen leicht eine Übersteuerung des Aufnahmepegels auftreten kann, zum anderen durch den Atmen des Sprechers zusätzlich starke akustische Störgeräusche (Windgeräu-

sche) entstehen können. Grundsätzlich wirkt sich auch ein hoher Geräuschpegel aus der Umgebung sehr störend auf die Erkennungsgenauigkeit des Spracheingabesystems aus.

- 5 Um diese bekannten Probleme zu lösen, werden beispielsweise für Computersysteme mit Texteingabe Mikrofonbügel eingesetzt, die (zumeist in Verbindung mit einem Kopfhörer oder Ohrhörer) vom Anwender am Kopf getragen werden. Bei entsprechender Ausrichtung des Bügels befindet sich das Aufnahmemikrofon auch bei Kopfbewegungen stets in gleichem Abstand nahe am Mund des Sprechers. Nachteilig ist dabei die eingeschränkte Bewegungsfreiheit aufgrund der üblichen Kabelverbindung zum Computer. Auch können Störgeräusche durch Bewegungen des Kabels auftreten. Zudem empfinden viele Personen das Tragen von Kopf- oder
- 10 Ohrhörern, insbesondere bei längerer Anwendung, als unangenehm.
- 15

- Alternativ werden daher auch Mikrofone mit stationärer Position, z.B. Tischmikrofon mit Stativ oder integriert in das Gehäuse (PC, Laptop) oder befestigt am Gerät (Türrahmen bei Sicherungsfunktion) verwendet. Nachteilig ist dabei die auf einen bestimmten Raumbereich vor dem Mikrofon beschränkte Aufnahmezone. Dies erfordert vom Anwender die Einhaltung einer definierten Position, Körperhaltung, Sprechrichtung usw.,
- 20 d.h. es besteht praktisch keine Bewegungsfreiheit während der Spracheingabe.
- 25

- Ausgehend von diesem Stand der Technik liegt der vorliegenden Erfindung die Aufgabe zugrunde, ein verbessertes System zur Spracheingabe zu entwickeln, das die genannten Nachteile weitgehend überwindet und zusätzliche Vorteile aufweist.
- 30

Diese Aufgabe wird bei einer Vorrichtung mit den Merkmalen des Oberbegriffs des Anspruchs 1 gelöst durch die kennzeichnenden Merkmale des Anspruchs 1. Weitere Einzelheiten der Erfindung und Vorzüge verschiedener Ausführungsformen sind Gegenstand der Merkmale der Unteransprüche.

Die erfindungsgemäße Vorrichtung wird im folgenden anhand einer bevorzugten Ausführungsform beschrieben, wobei Bezug genommen wird auf die Abbildungen und den darin aufgeführten Bezugsziffern.

Dabei zeigen:

Fig. 1 das erfindungsgemäße Spracheingabesystem aus
15 Zentraleinheit und separaten Sprachinterface

Fig. 2 eine schematische Darstellung einer Ausführungsform
des Sprachinterfaces

20 Fig. 3 Richtcharakteristik des vom Anwender mitgeführten
Sprachinterfaces

Die vorliegende Erfindung zielt darauf ab, Komfort und Qualität bei Spracheingabesystemen deutlich zu verbessern. Dabei soll ein mit Mikrofonen ausgestattetes transportables Sprachinterface vom Anwender ständig getragen werden, wodurch ihm letztlich ein universeller Sprachzugang zu verschiedenen Systemen ermöglicht wird. Durch Verwendung von Mikrofonarrays kann auch bei verschiedenen akustischen Umgebungen eine hohe Eingabequalität gegenüber Störgeräuschen erreicht werden. Damit ist es auch als Spracheingabesystem in Fahrzeugen geeignet, da Störungen durch Fahrgeräusche oder Echoeffekte bei

Lautsprecherausgaben vom Mikrofonarray vermindert werden. Für ein ständig zu tragenden Sprachinterfaces ist es wichtig, dass es klein und leicht ist und - je nach äußerer Gestaltung - beispielsweise als Schmuck oder Identifikationssymbol akzeptiert wird.

Fig. 1 zeigt als Übersicht die zusammenwirkenden Komponenten des Spracheingabesystems. Das Sprachinterface (2) ist als mobile Einheit ausgeführt und wird vom Anwender, z.B. an der Kleidung, getragen. Es überträgt die akustisch aufgenommenen Sprachsignale drahtlos, z.B. per Infrarot- oder Funkverbindung an die Zentraleinheit (1), in der die Signale weiter aufbereitet werden und diverse Steuerfunktionen ausgelöst werden.

15

Zur Gewährleistung einer hohen Qualität der Sprachaufnahme weist das Sprachinterface (2) zwei oder mehrere Mikrofone (3a, 3b, 3c) auf. Eine solche Anordnung ist in Fig. 2 vergrößert dargestellt.

20

Die verwendeten Mikrofone (3a, 3b, 3c) können individuelle Richtcharakteristiken (Niere, Hyperniere, Acht) aufweisen. Damit wird bereits durch eine vordefinierte Mikrofon-Richtcharakteristik Schall vorrangig aus einer bestimmten Zone verstärkt aufgenommen.

Die hier erfindungsgemäß vorgeschlagene Verwendung eines kleinen Mikrofonsystems mit zwei oder mehr Mikrofonen erlaubt die Bildung von Mikrofonarrays. Durch das Zusammenwirken der Mikrofone in einem solchen Mikrofonarray kann - in Verbindung mit einer für solche Arrays üblichen elektronischen Aufbereitung - die Qualität der Spracheingabe wesentlich verbessert

werden: So kann - über die oben erwähnte, unveränderliche Mikrofon-Richtcharakteristik hinaus - eine spezielle räumliche Richtwirkung des Mikrofonarrays eingestellt werden, d.h. akustische Signale werden vorrangig aus einem gewünschten

5 Raumbereich (Mundpartie des Anwenders) aufgenommen. Durch diese zusätzliche Array-Richtcharakteristik werden Nebengeräusche aus anderen Bereichen der Umgebung weiter unterdrückt bzw. können elektronisch weitgehend herausgefiltert werden.

10 Die Array-Richtcharakteristik ergibt sich aus der Anzahl und geometrischer Anordnung der Mikrofone. Im einfachsten Fall werden zwei Mikrofone verwendet (Minimalkonfiguration). Vorfürzugsweise ist das Interface jedoch mit drei (wie in Fig. 2 dargestellt) oder mehr Mikrofonen ausgestattet, die eine bes-

15 sere Richtwirkung und Störschallunterdrückung ermöglichen. Es gibt zwei grundsätzliche Mikrofon-Array-Anordnungen: 'broad-side' und 'end-fire'. Bei 'broad-side' liegt die Richtwirkung senkrecht auf der gedachten Verbindungsgeraden der Mikrofone, bei 'end-fire' liegt die Richtwirkung in Richtung der gedachte-

20 ten Verbindungsgeraden der Mikrofone. Das Ausgangssignal eines 'broad-side' Arrays ergibt sich in einfachster Form aus der Summation der einzelnen Signale, beim 'end-fire' Array aus der Differenz, wobei auch Laufzeitkorrekturen eingesetzt werden.

25

Die Richtwirkung des Mikrofonarrays kann durch weitere Maßnahmen verändert und damit eine adaptive Richtcharakteristik erreicht werden. Dabei werden die einzelnen Mikrofonsignale nicht einfach addiert oder subtrahiert, sondern durch spe-

30 zielle Signalverarbeitungsalgorithmen so ausgewertet, dass akustische Signale aus einer Hauptrichtung verstärkt empfangen werden und Nebengeräusche aus anderen Richtungen abgeschwächt

aufgenommen werden. Die Lage der Hauptrichtung ist dabei veränderbar, d.h. sie kann adaptiv an ein sich änderndes akustisches Szenario angepaßt werden. Dabei kann beispielsweise die Art und Weise, wie das Signal der Hauptrichtung bewertet und 5 maximiert wird und gleichzeitig Störgeräusche aus Nebenrichtungen minimiert werden, in einem Fehlerkriterium festgelegt sein. Algorithmen zur Erzeugung einer adaptiven Richtwirkung sind unter dem Begriff 'Beamforming' bekannt. Verbreitete Algorithmen sind z.B. der Jim-Griffith Beamformer oder der 10 'Frost' Beamformer.

Durch Veränderung entsprechender Parameter kann die Hauprichtung beim 'Beamforming' so variiert werden, dass sie mit der Richtung aus der gesprochen wird übereinstimmt, was einer 15 aktiven Sprecherortung entspricht. Eine einfache Möglichkeit die Sprechrichtung zu bestimmen ist z.B. die Laufzeitschätzung zwischen zwei Empfangssignalen zweier Mikrofone. Wird die Kreuzkorrelation zwischen den beiden Werten errechnet, so ergibt sich der maximale Kreuzkorrelations-Wert bei der Laufzeitverschiebung beider Signale. Wird nun die diese Laufzeit 20 zur Verzögerung des einen Signals eingesetzt, so sind die beiden Signale wieder gleichphasig. Damit ist die Hauptrichtung in die aktuelle Sprechrichtung eingestellt. Wird die Laufzeitschätzung und Korrektur wiederholt durchgeführt, so 25 wird einer Relativbewegung des Sprechers adaptiv gefolgt. Dabei ist es vorteilhaft für die Sprecherortung nur einen zuvor festgelegten Raumsektor zuzulassen. Dazu ist es erforderlich, die Mikrofonanordnung grob in einer gewissen Richtung zum Mund des Sprechers anzubringen, z.B. am Kleidungsstück des 30 Sprechers.

Der Mund des Anwenders kann sich dann innerhalb des festgelegten Raumsektors gegenüber der Position des Sprachinterfaces (2) frei bewegen, die Sprecherortung wird ihm folgen. Wird eine Signalquelle außerhalb des festgelegten Raumsektors

- 5 detektiert, dann wird diese als Störung identifiziert (beispielsweise eine Lautsprecherausgabe). Der Beamforming-Algorithmus kann sich nun speziell auf den Schall aus dieser Richtung im Sinne einer Minimierung des Störsignals einstellen. Dies lässt auch eine wirksame Echokompensation zu.

10

Fig. 3 zeigt beispielhaft ein kleines Mikrofonsystem aus zwei einzelnen Mikrofonen mit einer eingeprägten Richtcharakteristik die rechts am Mund des Sprechers vorbei gerichtet ist.

Die Mikrofone sind hier an der oberen Kante eines kleinen Gehäuses angebracht. Der Array-Typ ist 'broad-side', d.h. die Richtwirkung des Arrays ist senkrecht zur Kante des Gehäuses nach oben gerichtet. Die adaptive Richtwirkung über Beamforming-Algorithmen sorgt dafür, dass die wirksame Richtcharakteristik auf die Quelle, den Mund des Anwenders ausgerichtet ist.

20

Mikrofone mit hoher Qualität sind als Miniaturausführung (bis herab zu Millimetergröße) verfügbar. Ebenso sind drahtlose Übertragungseinrichtungen, z.B. Infrarot- oder Funksender, mit heutiger Technik (z.B. als SMD oder IC) extrem klein herstellbar. Zur Stromversorgung genügt eine kleine Batterie bzw. Akkumulator (z.B. Knopfzelle), da nur ein geringer Energieverbrauch entsteht. Damit ist die Integration aller Komponenten des Sprachinterfaces (2) zu einer kleinen und auch aufgrund des sehr geringen Gewichts bequem tragbaren Einheit möglich. Beispielsweise kann ein solches miniaturisiertes Sprachinterface (2) als Ansteckvorrichtung bzw. Clip an der

Kleidung befestigt werden (vergleichbar mit einer Brosche) oder an einem Armband oder einer Halskette mitgeführt werden.

In einer ersten Ausführungsform werden die jeweiligen Signale
5 der einzelnen Mikrofone des Arrays parallel zur Zentraleinheit übermittelt. Dort werden dann diese Signale elektronisch weiter verarbeitet, um Richtungscharakteristik und Störgeräuschunterdrückung einzustellen. Diese Funktionen können alternativ - zumindest teilweise - auch bereits im Sprachinterface selbst erfolgen. In dem Sprachinterface (2) sind dann
10 entsprechende elektronische Schaltkreise integriert, die für eine erste Signalaufbereitung sorgen. So kann beispielsweise auch der jeweilige Mikrofon-Aufnahmepiegel durch automatische Verstärkungsregelungen (Automatic-Gain-Control) angepaßt werden oder durch entsprechende Filter bestimmte Frequenzanteile
15 abgeschwächt oder verstärkt werden.

Wie in Figur 2 dargestellt kann das Eingabeinterface (2) auch Komponenten für eine Umwandlung analoger in digitale Signale
20 aufweisen. Die Übertragung digitaler Signale zur Zentraleinheit (1) bietet mit den heute verfügbaren Sende/Empfangstechniken einen sehr hohen Informationsfluss bei minimaler Störanfälligkeit.

25 Das Sprachinterface (2) kann mit einer Aktivierungstaste oder Sensorfläche ausgestattet werden, deren Betätigung/Berührung eine Spracherkennung aktiviert, beispielsweise durch Übertragung eines entsprechenden Signals an die Zentraleinheit (1). Alternativ kann eine solche Aktivierung auch durch Spracheingabe eines Schlüsselwortes erfolgen.
30

- In einer erweiterten Ausführungsform weist das Sprachinterface (2) zusätzlich eine Empfangseinrichtung auf (nicht dargestellt), die drahtlos übermittelte Steuersignale von der Zentraleinheit (1) empfangen kann. Damit kann die Charakteristik der Sprachaufnahme (Kombination der Mikrofone zum Array, Verstärkung, Frequenzfilter usw.) im Sprachinterface (2) durch die Zentraleinheit (1) verändert werden. Weiterhin kann die Zentraleinheit (1) über die Empfangseinrichtung dem Benutzer auch ein akustisches oder optisches Signal übermitteln, beispielsweise als 'Feedback' für erfolgreich ausgelöste Aktionen aber auch Fehlfunktionen (Kommando nicht erkannt oder - z.B. wegen Störung der Schnittstellen - nicht ausführbare Funktion) oder zur Anforderung weiterer Eingaben (Kommando unvollständig). Als Signalgeber können z.B. verschiedenenfarbige LED's oder piezokeramische Kleinlautsprecher in das Spracheingabeinterface (2) integriert werden. Auch ist es denkbar, eine Eingabeeinheit, beispielsweise Tastatur, zur Eingabe von Informationen in Textform vorzusehen.
- Die Zentraleinheit (1) empfängt die drahtlos übermittelten Signale des Sprachinterfaces (2) und wertet sie für eine Spracherkennung aus. Auch die verschiedenen erforderlichen Kalkulationen zur Sprecherortung und adaptiven Anpassung der Richtungscharakteristik des Mikrofonarrays können ganz oder teilweise in der Zentraleinheit (1) erfolgen. Die für eine zuverlässige und schnelle Signalverarbeitung und Spracherkennung erforderliche Prozessorleistung kann beispielsweise über ein übliches Computer- bzw. Mikroprozessorsystem zur Verfügung gestellt werden, das vorzugsweise Teil der Zentraleinheit (1) ist. Damit können anwenderspezifische Konfigurationen (Arraycharakteristik, Spracheigenheiten, spezielle

Schalt/Steuerkommandos usw.) jederzeit neu einprogrammiert werden.

Da die Zentraleinheit (1) getrennt vom Sprachinterface (2) -
5 z.B. integriert in Geräte (TV, Telefonanlage, PC-System usw.) - stationär ausgebildet ist, entstehen auch bei leistungsfähiger elektronischer und EDV-technischer Ausstattung keine besonderen Probleme oder Einschränkungen bzgl. Stromversorgung, Volumen, Gewicht oder auch Kühlung wärmeproduzierender Komponenten (Prozessor).

10 Die Zentraleinheit (1) ist so konfiguriert, dass erkannte Kommandos über integrierte Schnittstellen (4a, 4b, 4c) entsprechende Schalt- und Steuerfunktionen auslösen, auf die 15 verschiedene externe Geräte reagieren.

Hierzu können beispielsweise Telefonsysteme, Audio- und Videogeräte gehören, aber auch eine Vielzahl elektrisch/elektronisch steuerbarer Hausgeräte (Licht, Heizung, Klimaanlage, 20 Jalousien, Türöffner u.v.a.m). Bei Einsatz in einem Fahrzeug können entsprechend Funktionen des Fahrzeugs (Navigationssystem, Musikanlage, Klima, Scheinwerfer, Scheibenwischer usw.) gesteuert werden.

25 Die Übertragung der Steuersignale der Schnittstellen (4a, 4b, 4c) zu den verschiedenen externen Geräten kann ihrerseits wieder drahtlos (IR, Funk) oder auch (z.B. bei Einsatz im Fahrzeug) mittels Kabelverbindung erfolgen.

30 Die erfindungsgemäße Spracheingabevorrichtung bietet dem Anwender völlige Bewegungsfreiheit (auch über größere Distanz z.B. verschiedene Räume/Etagen eines Hauses). Hinzu kommt ein

optimaler Tragekomfort, da das Sprachinterface (2) nicht am Kopf befestigt werden muß. Durch das kleine Eigengewicht und die geringe Größe (insbesondere als Miniaturausführung) entsteht durch das Sprachinterface keinerlei Behinderung, so 5 dass es - vergleichbar zu einer Armbanduhr - für ein dauerhaftes Tragen geeignet ist.

Verschiedene Geräte bzw. Funktionen an unterschiedlichen Orten können mit ein und demselben Sprachinterface gesteuert 10 werden (Multimediaeinrichtungen, Hausgeräte, Fahrzeugfunktionen, Büro-PC, Sicherheitsanwendungen, usw.), da auch mehrere Zentraleinheiten an unterschiedlichen Standorten (z.B. Privaträume, Fahrzeug, Arbeitsbereich) so eingestellt werden können, dass sie auf das gleiche Sprachinterface bzw. den 15 gleichen Anwender reagieren.

Besonders vorteilhaft ist die hohe Flexibilität des vorgestellten Spracheingabesystems. So ermöglicht das Mikrofonarray eine aktive - je nach Ausführung auch automatische - Justierung auf veränderte Gegebenheiten, wie z.B. andere Raumakustik oder Befestigung des Sprachinterfaces an anderer 20 Stelle der Kleidung (andere Sprechrichtung) oder auch ein neuer Anwender (eigene Sprechweise). Durch Programmierung der Zentraleinheit kann das Spracheingabesystem flexibel an die 25 jeweiligen Steuer und Regelungsaufgaben angepaßt werden.

Grundsätzlich bietet das erfindungsgemäße System den Vorteil, dass das sehr kleine und leichte Sprachinterface vom Anwender ständig getragen werden kann. Ein Benutzer kann es ohne Be- 30 hinderung auch bei sich führen, wenn er seinen Aufenthaltsort ändert, z.B. in sein Auto einsteigt. Durch das Mikrofonarray und damit verbundenen Richtungswirkung der Schallaufnahme er-

gibt sich auch bei ungünstigen Umständen (Umgebungsgeräusche) eine hohe akustische Sprachaufnahmegeräuschequalität. Die aufwändigen Prozesse der Sprachaufbereitung können über entsprechend leistungsfähige Komponenten in der Zentraleinheit mit sehr 5 hoher Erkennungssicherheit erfolgen. Dabei kann die Spracherkennung individuell auf einen oder mehrere Anwender justiert werden, z.B. für eine Funktionsbeschränkung nur auf vorbestimmte, autorisierte Person(en). Die Ansteuerung verschiedener Geräte über entsprechende Schnittstellen ist frei konfigurierbar und kann nach individuellen Wünschen für verschiedene Anwendungszwecke modifiziert und erweitert werden.

DaimlerChrysler AG

Die/Fi

Patentansprüche

- 5 1. Spracheingabesystem aus stationärer Zentraleinheit (1) und transportablem Sprachinterface (2), wobei die Zentraleinheit (1) Mittel zum Empfang drahtlos übermittelter Sprachsignale und Mittel für eine Auswertung dieser Sprachsignale zur Erkennung gesprochener Worte aufweist, sowie Schnittstellen (4a, 4b, 4c), die in Abhängigkeit von erkannten Worten Schalt- bzw. Regelsignale erzeugen und an externe Geräte übertragen und wobei das Sprachinterface (2) zwei oder mehr Mikrofone (3a, 3b, 3c) aufweist, die als Mikrofonarray kombiniert sind, sowie Mittel zur drahtlosen Übertragung der vom Mikrofonarray generierten Signale an die Zentraleinheit (1).
- 10 2. Spracheingabesystem nach Anspruch 1, dadurch gekennzeichnet, dass das Sprachinterface (2) Mittel zur elektronischen Aufbereitung der von den Mikrofonen (3a, 3b, 3c) des Mikrofonarrays erzeugten Signale aufweist.
- 15 20 3. Spracheingabsystem nach Anspruch 2, dadurch gekennzeichnet, dass die elektronische Aufbereitung eine Veränderung der jeweiligen Signalpegel (Verstärkung, Dämpfung, AGC) und/

oder Frequenzanteile (Höhen-, Tiefen-, Band-Filter) und/
oder der jeweiligen Signalphasen bewirkt.

4. Spracheingabesystem nach einem der Ansprüche 2 oder 3,
5 dadurch gekennzeichnet,
dass die Mittel zur elektronischen Aufbereitung auch Mit-
tel zur Umwandlung analoger in digitale Signale beinhal-
ten.

10 5. Spracheingabesystem nach einem der Ansprüche 1 bis 4,
dadurch gekennzeichnet,
dass das Sprachinterface (2) Mittel zum Empfang drahtlos
(z.B. IR, Funk) von der Zentraleinheit (1) übermittelter
Signale aufweist.

15 6. Spracheingabesystem nach Anspruch 2, 3 oder 4,
dadurch gekennzeichnet,
dass das Sprachinterface (2) Mittel zum Empfang drahtlos
(z.B. IR, Funk) von der Zentraleinheit (1) übermittelter
20 Steuersignale aufweist, mit denen die Zentraleinheit (1)
die Mittel zur elektronischen Aufbereitung des Sprachin-
terfaces (2) beeinflussen kann.

25 7. Spracheingabesystem nach einem der Ansprüche 1 bis 6,
dadurch gekennzeichnet,
dass die Mittel der Zentraleinheit (1) zur Auswertung der
Sprachsignale als Prozessorsystem (EDV) ausgeführt sind.

30 8. Spracheingabesystem nach Anspruch 7,
dadurch gekennzeichnet,
dass das Prozessorsystem Anwender-bezogene Daten (z.B.
individuelle Sprachkommandos, Sprachparameter) speichert

und zur Auswertung der Sprachsignale und/oder zur Erzeugung individuell zugeordneter Steuersignale einbezieht.

9. Spracheingabesystem nach einem der Ansprüche 1 bis 8,
5 durch gekennzeichnet,
dass das Mikrofon-Array (3a, 3b, 3c) eine Richtcharakteristik in der Weise aufweist, dass Schall vorrangig aus einer Richtung aufgenommen wird.

10 10. Spracheingabesystem nach Anspruch 9,
durch gekennzeichnet,
dass die Richtcharakteristik durch Aufbereitung der von
den Mikrofonen (3a, 3b, 3c) des Mikrofonarrays erzeugten
Signale über Beamforming-Algorithmen (z.B. Griffith-,
15 Frost-Beamformer) verändert werden kann.

11. Spracheingabesystem nach Anspruch 10,
durch gekennzeichnet,
dass die Richtcharakteristik in der Weise verändert wird,
20 dass sie einer räumlichen Verlagerung der Schallquelle
folgt (Sprechernachführung).

12. Spracheingabesystem nach einem der Ansprüche 1 bis 11
durch gekennzeichnet,
25 dass die einzelnen Mikrofone (3a, 3b, 3c) des Arrays eine
Richtcharakteristik (z.B. Niere, Acht) aufweisen.

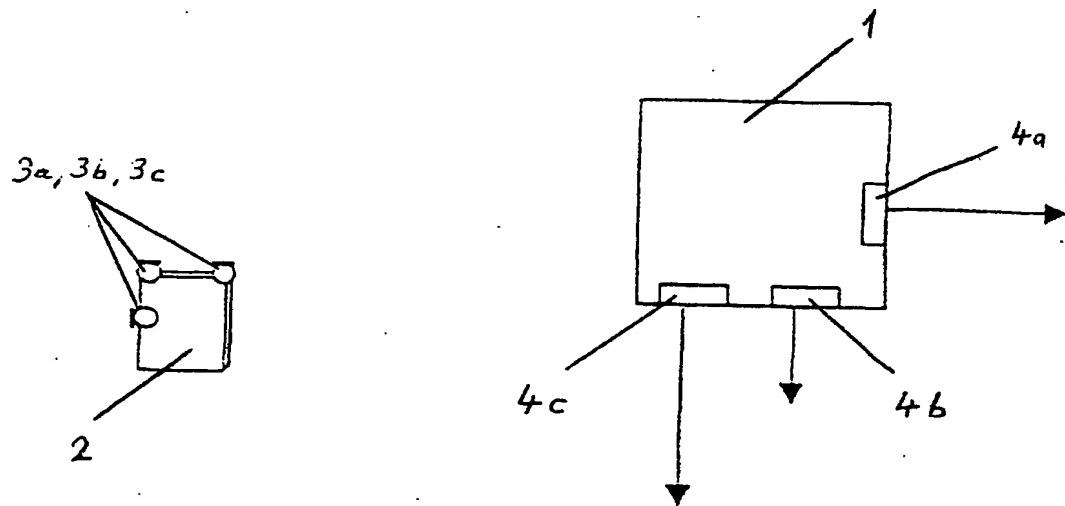


Fig. 1

5

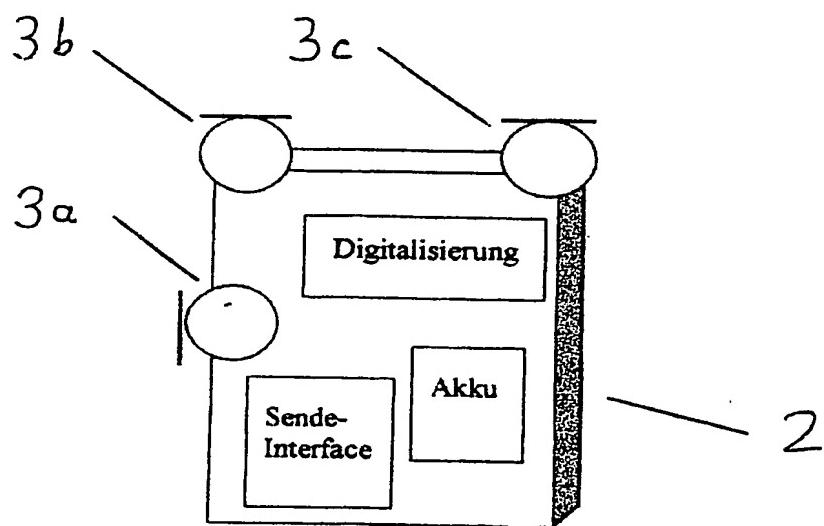
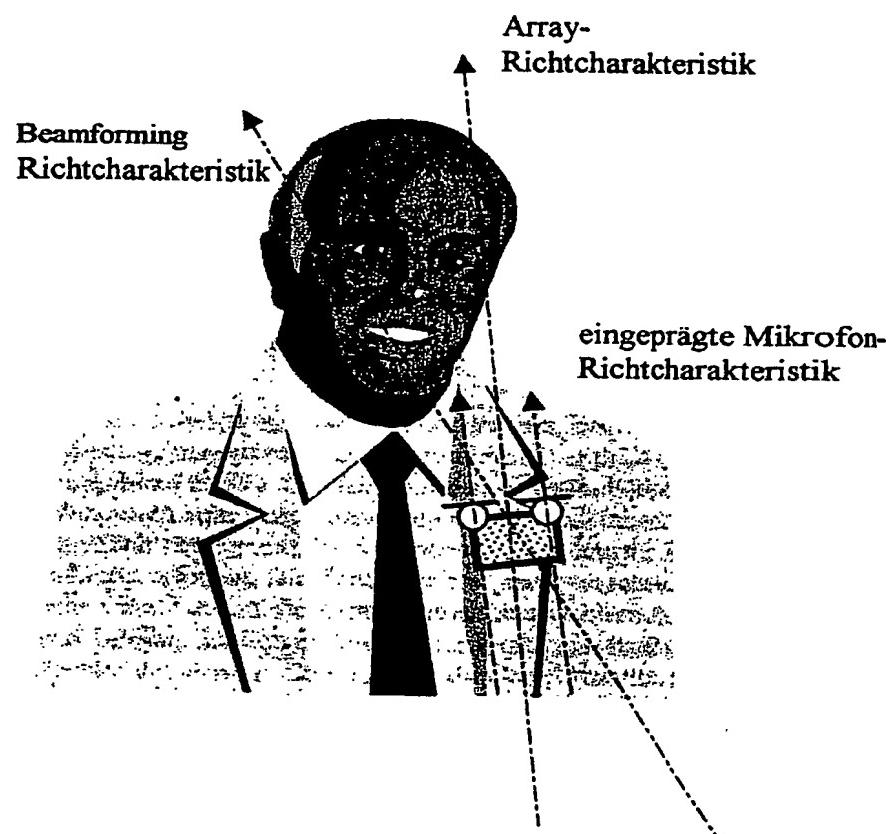


Fig. 2

5



10

Fig. 3

DaimlerChrysler AG

Die/Fi

ZusammenfassungSpracheingabe-Interface

5 Beschrieben wird ein Spracheingabesystem, bestehend aus einer stationären Zentraleinheit (1) und einem transportablen, vom Anwender z.B. an der Kleidung mitgeführten Sprachinterface (2). Das Sprachinterface (2) dient zur akustischen Aufnahme gesprochener Kommandos, die nach elektronischer Aufbereitung drahtlos an die Zentraleinheit (1) übertragen werden. Dabei enthält das Sprachinterface zwei oder mehr Mikrofone (3a, 3b, 3c), die als Mikrofonarray zusammenwirken, wodurch einerseits eine Richtcharakteristik bei der Schallaufnahme und andererseits eine weitgehende Unterdrückung von Störgeräuschen ermöglicht wird. Die Zentraleinheit (1) enthält die Komponenten der Spracherkennung, z.B. Prozessorsystem, und Schnittstellen (4a, 4b, 4c), über die externe Geräte drahtlos oder über Kabel gesteuert werden.

10

15

20

(Fig. 2)

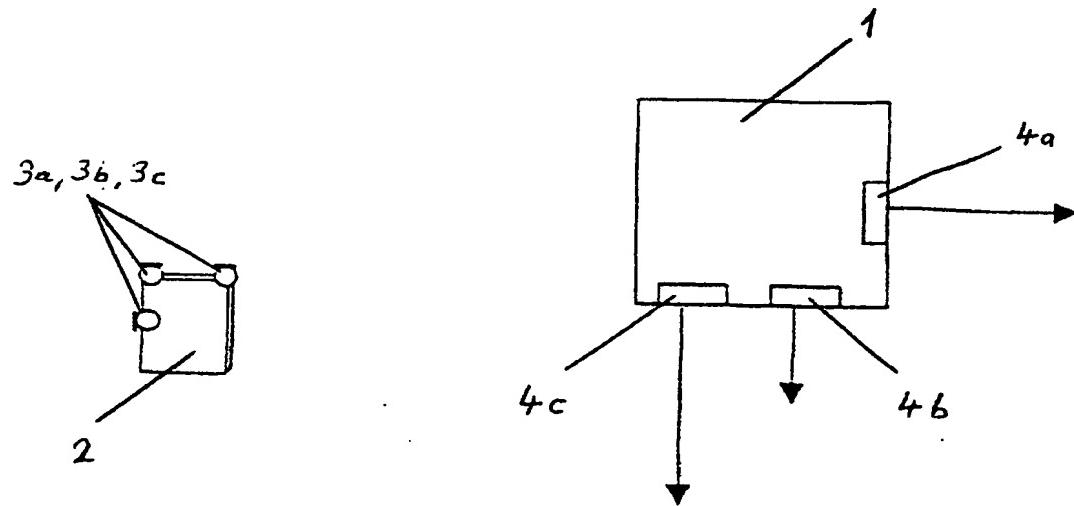


Fig. 1

5

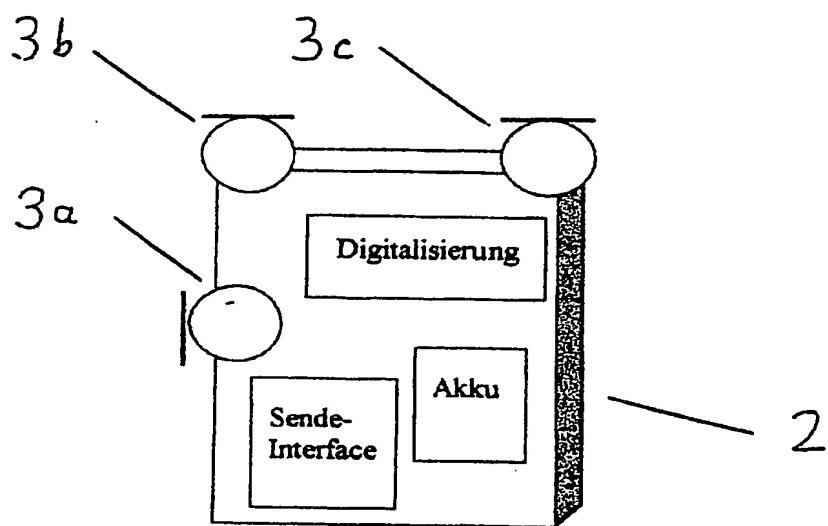


Fig. 2

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record.**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.